San Francisco Bay Region Historical Earthquake Re-analysis Project

Grant No. 00-HQ-GR-0044

Robert Uhrhammer

Berkeley Seismological Laboratory, 207 McCone Hall #4760
University of California, Berkeley, CA 94720-4760
Telephone: 1.510.642.8504
Fax: 1.510.643.5811
E-mail: bob@seismo.berkeley.edu

Program Element: I
Keywords: Earthquake Probabilities, Source Characteristics,
Probabilistic Seismic Hazard, Database

FINAL REPORT

INVESTIGATIONS UNDERTAKEN:

Introduction

The objective is to characterize the spatial and temporal evolution of the San Francisco Bay Region (SFBR) seismicity during the initial part of the earthquake cycle as region emerges from the stress shadow of the great 1906 San Francisco earthquake. The problem is that the existing Berkeley Seismological Laboratory (BSL) seismicity catalog for the SFBR, which spans most of the past century (1910-present), is inherently inhomogeneous because the location and magnitude determination methodologies have changed, as seismic instrumentation and computational capabilities have improved over time.

Creation of a SFBR catalog of seismicity that is homogeneous, that spans as many years as possible, and that includes formal estimates of the uncertainty in the estimated parameters is a fundamental prerequisite for probabilistic studies of SFBR seismicity. The existence of the invaluable BSL seismological archive, containing the original seismograms as well as the original reading/analysis sheets, coupled with the BSL computational capabilities allows the application of modern analytical algorithms towards the problem of determining the source parameters, including formal uncertainties, of the historical SFBR earthquakes.

Our approach is to systematically re-analyze the data acquired from the archive to develop a homogeneous SFBR catalog of earthquake location and local magnitude (M_L), including formal uncertainties on all parameters which extends as far back in time as the instrumental records allow and which is complete above appropriate threshold magnitudes.

Background

Although the 1910 to present BSL catalog of earthquakes for the SFBR appears to be a simple list of events, one must remember that it really is a very complex data set. It is easy to misinterpret observed variations in seismicity if one does not understand the limitations of this catalog. The existing 1910 to present BSL catalog of earthquakes for the SFBR is inhomogeneous in that it suffers from the three types of man-made seismicity changes identified by Habermann [1987], namely detection changes, reporting changes, and magnitude shifts. The largest change in the detection capability of the BSL seismic station network occurred circa 1927-1931 with the installation of Wood-Anderson seismometers at four SFBR stations (BRK, MHC, PAC, and USF) and the resulting increase in sensitivity lowered the threshold for detection by about 1.5 M_L units. The most significant reporting changes occurred circa 1941-1942 when the BSL Bulletin entries changed from descriptive to geographical coordinate locations and the reporting of local magnitude began. A possible magnitude shift occurred in 1954 when the response of the Wood-Anderson seismographs was allegedly changed (owing to a reported changing the free period from 1.0 to 0.8 seconds) [Bolt and Miller, 1975]. No corroborating evidence for this shift has been found, however.

The lack of a homogeneous catalog of earthquakes for the SFBR which spans most of the past century, the availability of the invaluable BSL seismological archive, the interest in the Working Group on California Earthquake Probabilities [WGCEP 1999], the funding of an initial effort with support from the USGS/PG&E CRADA, and the purchase and loan of a high-resolution wide-format digitizer by the USGS, combine to provide both an incentive and an unique opportunity to systematically re-analyze, using modern algorithms, the BSL seismographic records and data for SFBR earthquakes and to produce a homogeneous catalog of earthquakes for the region.

Initial Effort

During the summer of 1998, the USGS, via the USGS/PG&E CRADA, funded two students to transcribe the data from the original BSL reading/analysis sheets to computer readable form. With this funding they were able to transcribe the reading/analysis for SFBR earthquakes, working back in time from 1983 to 1944 (1983 onward was already in computer database). The 1951-1998 data were used to systematically determine the location and local magnitude (including their formal uncertainties) of earthquakes which have occurred in the SFBR and an interim catalog of these events has been compiled and is available via URL: http://perry.geo.berkeley.edu/seismo/herp/. The interim catalog starts in 1951 because the maximum trace amplitudes, used in the determination of M_L, were not registered on the reading/analysis sheets prior to that time.

HERP Effort

Our HERP effort has concentrated on the processes involved with expanding the database of SFBR earthquakes. This involves: compiling a list of candidate SFBR events using a combination of the UCB online catalog and the events listings in the BSL Bulletins; reading the maximum trace amplitudes recorded on the Wood-Anderson seismograms (~1932 forward) and on the Bosch-Omori and Wiechert seismograms (pre-1932), and; transcription of the amplitude data and the reading/analysis sheet data. With the completion of this project, we now have a computer readable database of SFBR earthquake phase and amplitude data, complete at appropriate magnitude thresholds (M_L ~3 from ~1930 and M_L ~4.5 pre ~1930), which spans approximately the past 90 years.

Concurrent with the transcription and seismogram reading effort, we compiled a list of candidate events (M_L ~4.5 and larger) for eventual scanning and digitizing of the relevant seismograms for moment tensor analysis etc. This is important particularly for the pre-1964 SFBR events because the waveforms only exist on paper seismograms. From 1964 onward until the advent of digital recording at BSL, broadband seismic records that are also stored on analog magnetic tapes which can be digitized upon playback in lieu of scanning and digitizing the paper seismograms. The primary incentive for digitizing the analog tape recordings is that the resulting resolution is a factor of eight higher than the resolution which can be obtained from scanning and digitizing the paper records. The BSL analog magnetic tape archive contains short-period recordings from 8-11 BSL SFBR seismic stations (varied over time) and broadband recordings from three SFBR seismic stations. Any tape playback and tape digitization effort will need to be a joint effort between the BSL which has the tape archive and the USGS Menlo Park that has the playback and digitization hardware. At the present time, however, it is unclear as to when funding or manpower will be available for such a joint project tape digitization project.

Also concurrent with the above efforts, we are searching for candidate repeating SFBR earthquakes. Statistically, we expect to find repeating earthquakes in the SFBR with recurrence intervals of order 10 to 50 years and since the new catalog will span 70+ years at the M_L ~3 threshold, we could potentially identify repeating sequences containing up to 5 events, say. Once a sequence of repeating earthquakes is identified, there are a couple of intriguing possibilities. Once older earthquakes have been identified as members of a repeating sequence, we can get a much better estimate of their size and location by assuming that they are very nearly the same as their most recent sequence members. This offers a powerful and robust method for constraining the hypocenter location and magnitude of the older earthquake when constructing the SFBR earthquake catalog and it also provides a means of checking the spatial and magnitude homogeneity of the catalog.

ACCOMPLISHMENTS:

The reading of maximum trace amplitudes, registered on the Wood-Anderson seismograms for the 1927-1950 SFBR candidate events, and subsequent transcription to the database is complete. This was a very labor intensive and time consuming task because it involves retrieving, reading, and re-archiving of the original seismograms for four SFBR stations (BRK, MHC, PAC, and USF shown in Figure 1) that are kept in the BSL seismogram archive. While reading the Wood-Anderson seismograms, particularly for sequences of events that occurred circa 1942 through 1945, we identified several SFBR candidate events (mostly the smaller members of earthquake sequences) for which there was no data in either the Bulletin or in the reading/analysis sheets. These events were subsequently read and added to the database.

A list of SFBR candidate events for 1932-1941 has been compiled from the information in the BSL Bulletins. This task was complicated by the fact that there was rarely any local magnitude information available and that, prior to July 1941, the locations were descriptive (with geographical coordinates rarely provided) and the event times were only accurate to within a couple of minutes [Byerly and Meeker, 1948]. Consequently, most of the SFBR candidate events, which occurred prior to July 1941, are not in the UCB online catalog (available from the Northern California Earthquake Data Center via URL: http://quake.geo.berkeley.edu/ncedc/catalog-search.html). An algorithm was subsequently developed to convert the descriptive locations to geographical coordinates and P-wave and S-wave onset times (registered in the Bulletin) at generally the closest station were used to estimate the origin time to within a few seconds. The resulting list contained 721 candidate SFBR events (see Figure 1) while the UCB online catalog derived list contained only 270 events for the same area and time span. We read the Wood-Anderson seismograms and transcribed the resulting amplitude and phase data for the 1927-1941 SFBR candidate events.

Analyzing the 1910-1927 data has proven quite a challenge. There are only two stations reported in the BSL Bulletin (BRK and MHC) prior to 1927 and the available event information from the UCB online catalog is rather sparse with the origin time only accurate to a within a couple of minutes and the location typically given only in terms of distances from the stations and/or where the event was reported felt. To resolve ambiguity in the location when using data from only two stations, we resort to felt reports and comments from the original reading/analysis sheets and also, when the need arises, to the extensive newspaper archive in the UCB main library. Determination of the magnitude of these events is based on amplitudes registered on the Bosch-Omori and Wiechert seismographs (generally reported in microns on the reading/analysis sheets). We have identified 360 candidate SFBR events that occurred between 1910 and 1931 and we have transcribed the available data from the reading/analysis sheets.

As noted above, the pre-1960 SFBR events were generally recorded by the four stations (BRK, MHC, PAC and SFB; shown in Figure 1) located in the region and the phase data (typically P- and S-wave onset times) recorded on the reading/analysis sheets have a generally unknown level of accuracy owing to the presence of occasional clock errors, reading errors, and/or transcription errors (station SFB was particularly troublesome with frequent clock errors so only the S-P time interval was reliable). Also, the pre-1927 SFBR events were only recorded by two stations (BRK and MHC) which presents a challenge in that the resulting location ambiguity needs to be resolved. Location algorithms that employ the minimization of a misfit function do not produce robust solutions when these types of errors or problems are present in a sparse data set. In order to obviate the man-power intensive reading of phase onset times from the original seismograms, a fuzzy logic (Yen and Langari, 1999) based location algorithm (FL_RELP; Fuzzy Logic Regional Earthquake Location Program) for locating the SFBR earthquakes has been developed. The FL_RELP algorithm utilizes and builds upon some of the components that were developed for its predecessor, namely BW_RELP (Broadband Waveform Regional Earthquake Location Program) (Dreger et al., 1997; Uhrhammer et al., 2001). FL_RELP uses a adaptive grid search algorithm to map the possibility function for all observational subsets and also to determine the maximum in the cumulative possibility function. The inconsistent data are essentially inhibited from influencing the solution by using an appropriate threshold for mapping the possibility function. The most powerful argument in favor of the fuzzy logic approach is it ability to obtain a robust solution when using a sparse set of imprecise observed data that may contain inaccuracies and/or errors. Preliminary tests indicate that up to approximately one-third of the data can have large residuals (in the classical least-squares sense) without significantly influencing the solution. This capability along with the fact that the algorithm does not require the determination of which data are inaccurate or erroneous means that the algorithm is also ideally suited for use in automated seismic event location schemes.

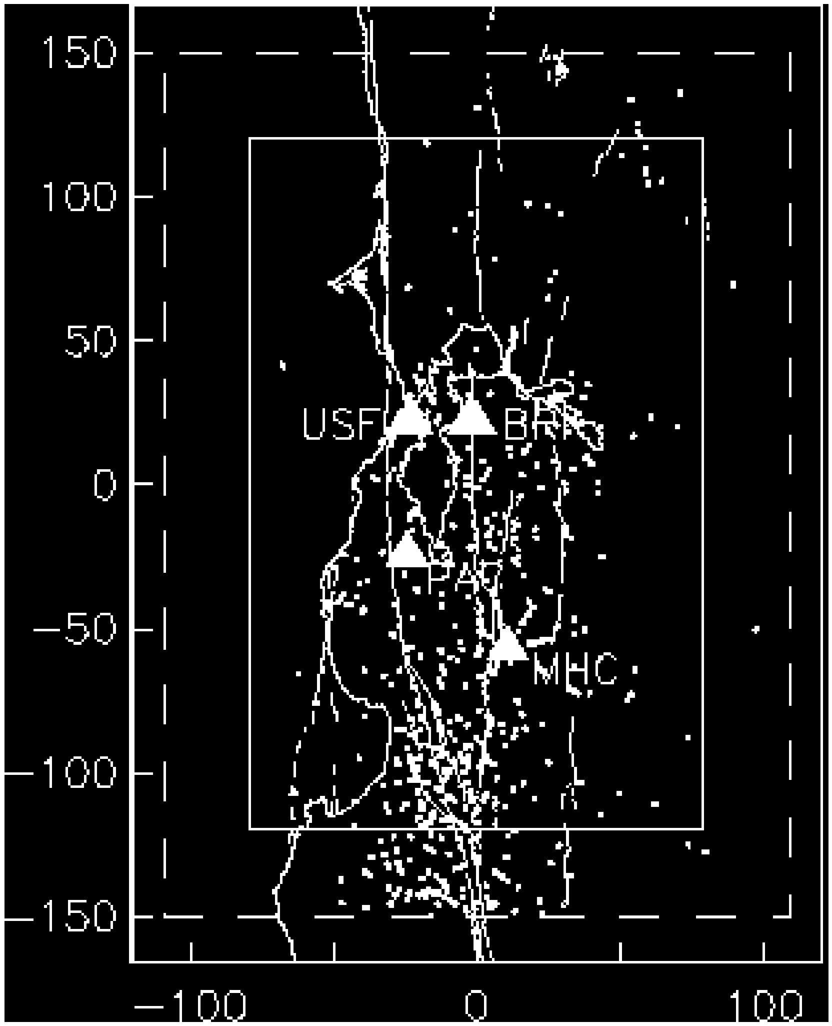

Figure 1. Map of the SFBR study area (inner box) and 30 km buffer zone (dashed box) showing the locations of the stations housing Wood-Anderson seismographs (triangles) and the preliminary locations of the 721 candidate SFBR earthquakes that occurred between 1932 and 1941 (dots). The distances are in km. The "candidate" SFBR events are those events which have an initial location estimate within the SFBR (including a 30 km buffer zone around the SFBR) and which either have no M_L assigned or they have an initial M_L estimate of either 2.8 or larger (1932-present) or 4.3 or larger (pre-1932). In order to minimize the number of seismograms in the BSL archive that need to be analyzed, we have been scanning the microfilm records of the seismograms from the Lick Observatory station to cull events from the SFBR candidate list which are smaller than the appropriate threshold M_L. The microfilming of the 1911 through June 1962 MHC seismograms was done in 1981 as a pilot effort of the Historical Seismogram Filming Project [Uhrhammer, 1983].

UNANTICIPATED PROBLEMS:

While we expected the reading and analysis of the seismograms to be time consuming, we found that it was taking more time than we anticipated owing largely to the fact that the seismogram archive, the reading/analysis sheet archive, and the BSL facilities with the computer work stations and the high-resolution scanner were in different buildings scattered across the Berkeley campus. The students working on the project had to spend time at three separate sites to do the work and this proved to be a somewhat inefficient use of their time. To expedite the processing, we adopted a modern factory model and reorganized one room in McCone Hall (the main BSL facility) for doing all the required reading, transcription, and analysis. We now transport the seismograms and the reading/analysis sheets to McCone Hall as they are needed and re-archive them when the processing is done (in annual batches). This change successfully increased our overall productivity in reading, transcribing, and analyzing the SFBR earthquakes.

In order to compile a list of candidate highly-similar (repeating) earthquakes we developed an algorithm to search for all sets of earthquakes which have nearly identical observed S-P time intervals observed at the SFBR stations, which have similar magnitudes, and which form a roughly periodic sequence in time. This candidate list greatly reduces the number of events that need to be examined to identify sequences of repeating earthquakes. The Wood-Anderson seismograms play a crucial role in this analysis since they recorded SFBR earthquakes over a time span of more than six decades. However, we found that we were effectively thwarted in this effort by the bandwidth limitation of the scanning hardware/digitizing software that we had available for processing the Wood-Anderson seismograms. The spatial resolution at 1600 dpi (the highest available on our scanner) effectively limited the frequency content of the resulting digitized seismogram trace to 2 Hz and this is insufficient to use correlation analysis to reliably identify highly-similar earthquakes.

DISCUSSION and CONCLUSIONS:

The HERP effort deals primarily with the transcription and analysis BSL pre-1984 data because the data from BSL data from 1984 to the present exists in computer readable form. The HERP effort can be basically divided into three time epochs: pre-Wood-Anderson recording (circa 1927-1932); pre-routine ML determination (circa 1950), and; pre-routine daily entry of the phase and amplitude data into a computer database (1984-present). Of these epochs, the 1950-1983 epoch required the least effort because the data existed on the original BSL reading/analysis sheets and the primary task was to transcribe the relevant data to computer readable flat files and to verify the accuracy of the transcription. The 1927-1950 data required a considerable amount of effort because the Wood-Anderson maximum trace amplitude data, which is required for the determination of local magnitude and its uncertainty, was not entered on the original BSL reading/analysis sheets. Consequently, we had to pull the original Wood-Anderson seismograms from the BSL Seismogram Archive, read and transcribe the maximum trace amplitudes for the events of interest, and then re-file the seismograms in the archive. This effort was successful and it expanded the number of pre-1950 SFBR events for which ML was determined by 44.3 percent from 329 to 475. Analysis of the pre-Wood-Anderson recording data, on the other hand, was problematical because there were only two BSL SFBR seismic stations in routine operation at that time (BRK and MHC) and the magnification of the seismographs was of order 100 (about 5% of the sensitivity of a Wood-Anderson). As a consequence, the threshold for magnitude determination increases to approximately 4.5.

Ultimately 2359 candidate SFBR earthquakes, which occurred between 1912 and 1983, were transcribed and locations and local magnitudes were determined for 1806 SFBR events. The resulting SFBR HERP seismicity catalog is available at URL: ftp://ncedc.org/pub/herp.

A composite 1932-2002 ML≥3 SFBR seismicity catalog was compiled to characterize the spatial and temporal evolution of the seismicity during the initial part of the earthquake cycle as region emerges from the stress shadow of the great 1906 San Francisco earthquake. The catalog was compiled by merging the HERP catalog (1932-1983) and the BSL BDSN catalog (1984-2002) (available via URL: http://www.ncedc.org/ncedc/catalog-search.html.). The resulting catalog contained 3163 earthquakes. To analyze temporal and spatial changes in the rate of seismicity, the catalog was first declustered and a SFBR catalog of 2034 mainshocks was compiled.

A plot of the mainshock annual rate of seismicity is shown in Figure 2. Note that there is a significant demarcation in the annual seismicity rate circa 1983 when the average annual rate of ML≥3 mainshocks increased markedly from less than 30 per year to more than 40 per year. Prior to circa 1983 the annual rate of seismicity had a roughly linear increasing trend and after that time the rate increased markedly with the occurrence of large earthquakes in the SFBR during the 1980s. This is evidence that the SFBR is emerging from the stress shadow of the great 1906 San Francisco earthquake.

To characterize the spatial evolution of the SFBR seismicity, the distrubution of annual spatial epicenter centroids was calculated from the SFBR mainshock catalog. The results are tare shown in Figure 3 and given the 41+ km standard error of each centroid, there is no evidence that the spatial distribution of the SFBR seismicity has changed significantly during the past 70+ years while the SFBR is emerging from the stress shadow of the great 1906 San Francisco earthquake.

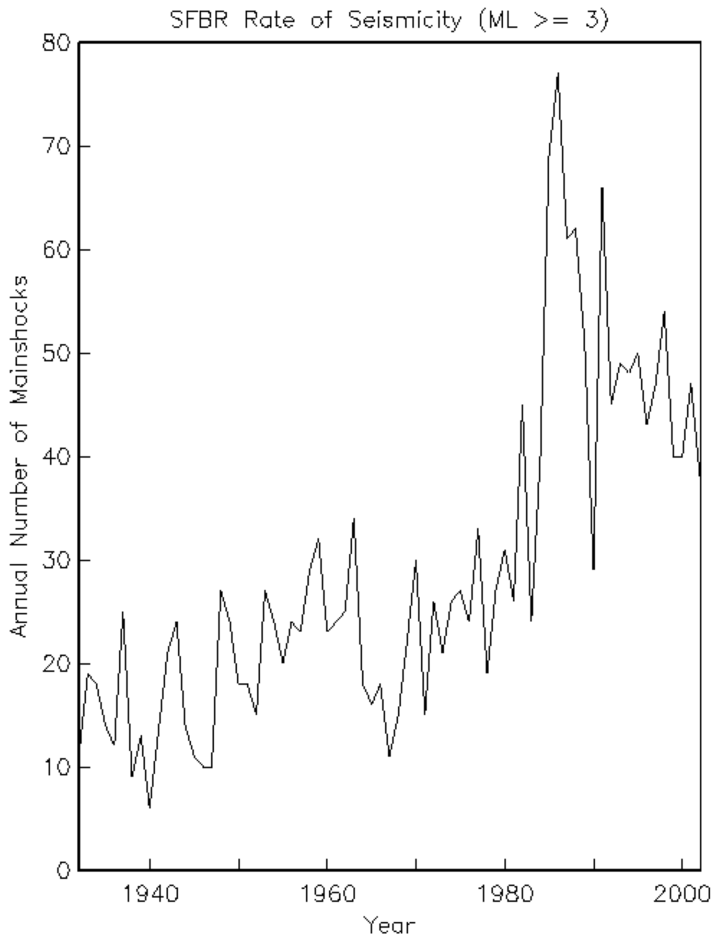

Figure 2. Annual rate of SFBR seismicity. The data have been declustered so that only the largest event in each sequence is included in the tally of earthquakes.

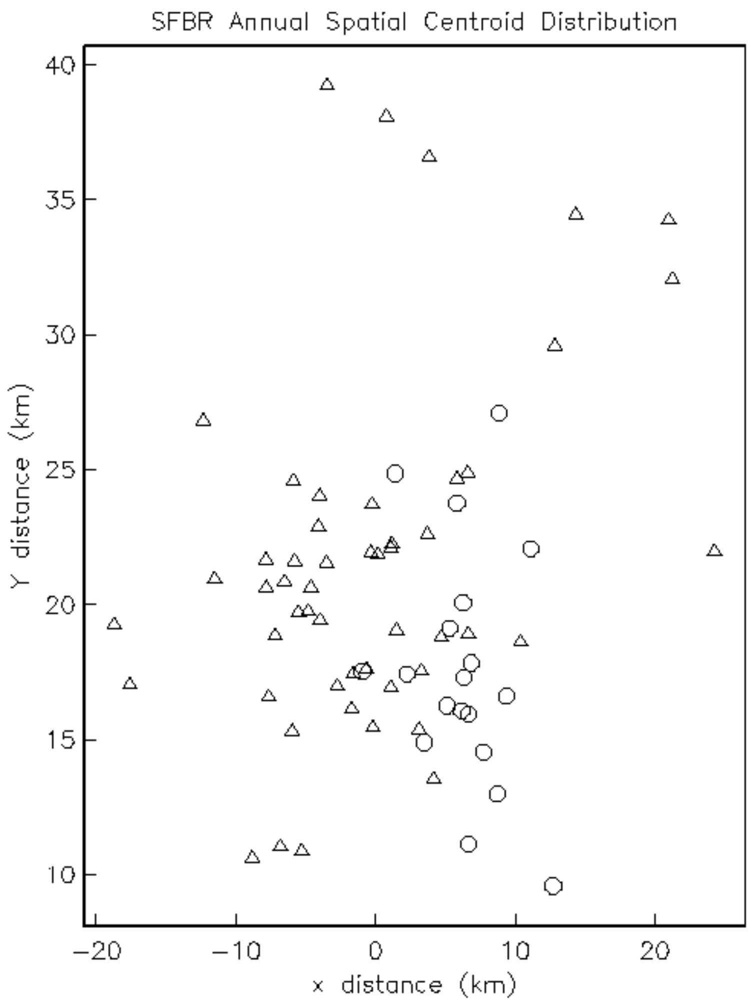

Figure 3. Spatial distribution of annual mainshock centroids plotted using the same coordinates as in Figure 1. The triangles are pre-1983 and the octagons are 1983-2002 data. The minimum centroid standard error is 41 km.

NON-TECHNICAL SUMMARY:

This project focuses on the creation of a seismicity catalog for San Francisco Bay Region earthquakes that is homogeneous, that spans as many years as permitted by the instrumental records kept on store in the Berkeley Seismological Laboratory archives and that includes formal estimates of the uncertainty in the estimated parameters. Such a catalog is a fundamental prerequisite for probabilistic studies of the regions seismicity and for characterizing the spatial and temporal evolution of the seismicity during the initial part of the earthquake cycle as the region emerges from the stress shadow of the great 1906 San Francisco earthquake. The data show that the rate of seismicity in the San Francisco Bay Region is increasing indicating that the region is indeed emerging from the stress shadow and also that there is no evidence that the spatial distribution of the earthquakes is changing significantly.

REPORTS PUBLISHED:

Uhrhammer, R. A., J. Fink, and S. Ford., San Francisco Bay Region Historical Earthquake Relocation Project, *Seism. Research Lett.*, 70, 271, 1999.

Uhrhammer, R. A., The San Francisco Bay Region Historical Earthquake Re-analysis Project: A progress report and locating pre-1960 instrumentally recorded earthquakes using a fuzzy-logic-based algorithm. Seism. Res. Lett., 72, 245, 2001.

Wesson, R. L., W. H. Bakun, R. A. Uhrhammer, D. H. Oppenheimer, and D. M. Perkins, Application of Baysian Inference to the Association of Earthquakes in the San Francisco Bay Region, *Seism. Research Lett.*, 70, 270, 1999.

Berkeley Seismological Laboratory Annual Report HERP Contributions URLs:
http://seismo.berkeley.edu/annual_report/ar98_99/node14.html
http://seismo.berkeley.edu/annual_report/ar99_00/node26.html
http://seismo.berkeley.edu/annual_report/ar00_01/node21.html
http://seismo.berkeley.edu/annual_report/ar01_02/node23.html
http://seismo.berkeley.edu/annual_report/ar02_03/node19.html

The data and analysis files are available on-line as discussed in the Data Availability section below.

References

Bolt, B. A. and R. D. Miller, Catalog of Earthquakes in Northern California and Adjoining Areas: 1 January 1910 - 31 December 1972, *Seismographic Stations, University of California, Berkeley*, iv+567 pp., 1975.

Byerly, P. and J. E. Meeker, Bulletin of the Seismographic Stations, *University of California Press*, 11, No. 2, 45-107, 1948.

Dreger, D., M. Pasyanos, J. Frank and B. Romanowicz, Evaluation of the Performance of Broadband Stations and Regional Arrays in Global Monitoring - Phase II Final Report, *submitted July 1997,* 10 pp.

Habermann, R. E., Man-made Changes of Seismicity Rates, *Bull. Seism. Soc. Am.*, 77, 141-159, 1987.

Romney, C. F. and J. E. Meeker, Bulletin of the Seismographic Stations, *University of California Press*, 18, pp. 89, 1949.

Uhrhammer, R. A., Microfilming of Historical Seismograms from the Mount Hamilton (Lick) Seismograph Station, *Bull. Seism. Soc. Am.*, 73, 1197-1202, 1983.

Uhrhammer, R. A, D. Dreger and B. Romanowicz, Best Practice in Earthquake Location Using Broadband Three-Component Seismic Waveform Data, *Pure Appl. Geophys., Vol. 158, no. 1/2, pp. 259-276,* 2001.

Uhrhammer, R. A., The San Francisco Bay Region Historical Earthquake Re-analysis Project: A progress report and locating pre-1960 instrumentally recorded earthquakes using a fuzzy-logic-based algorithm. Seism. Res. Lett., 72, 245, 2001.

WGCEP 1999, Earthquake Probabilities in the San Francisco Bay Region: 2000 to 2030 - A Summary of Findings, Working Group on California Earthquake Probabilities, USGS Open File Report 99-517, 1999, version 1.0, available online via URL: http://quake.wr.usgs.gov/study/wg99/of99-517/index.html.

Yen, J. and R. Langari, Fuzzy Logic: Intelligence, Control, and Information, 548 + xxi pp., Prentice Hall, New Jersey, 1999.

Data Availability

The transcribed HERP data files for 1912-1983 for the SFBR and the resulting analysis and seismicity catalog files are available online via URL: ftp://ncedc.org/pub/herp/. Information about the contents of each sub-directory is given in the associated README file in each directory. For further information, contact Robert Uhrhammer by e-mail (bob@seismo.berkeley.edu).